

A SYSTEM AND METHOD FOR RETRIEVING DOCUMENTS OR SUB-DOCUMENTS BASED ON EXAMPLES

ABSTRACT

5

Disclosed are a system, method, and program storage device implementing the method of extracting information, wherein the method comprises inputting a query; searching a database of documents based on the query; retrieving documents from the database matching the query using a plurality of classifiers arranged in a hierarchical cascade of
10 classifier layers, wherein each classifier comprises a set of weighted training data points comprising feature vectors representing any portion of a document; and weighing an output from the cascade according to a rate of success of query terms being matched by each layer of the cascade, wherein the weighing is performed using a terminal classifier.